Carbohydrate Research 434 (2016) 99–106

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Site-directed mutagenesis of bacterial cellulose synthase highlights sulfur–arene interaction as key to catalysis

Shi-jing Sun, Yoshiki Horikawa [1], Masahisa Wada [2, 3], Junji Sugiyama, Tomoya Imai*

Research Institute for Sustainable Humanosphere (RISH), Kyoto University, Uji, Kyoto 611-0011, Japan

ARTICLE INFO

Article history:
Received 22 June 2016
Received in revised form
26 August 2016
Accepted 27 August 2016
Available online 29 August 2016

Keywords:
Cellulose synthase
CesA
Site-directed mutagenesis
Gluconacetobacter xylinus
Sulfur–arene interaction

ABSTRACT

Cellulose is one of the most abundant biological polymers on Earth, and is synthesized by the cellulose synthase complex in cell membranes. Although many cellulose synthase genes have been identified over the past 25 years, functional studies of cellulose synthase using recombinant proteins have rarely been conducted. In this study, we conducted a functional analysis of cellulose synthase with site-directed mutagenesis, by using recombinant cellulose synthase reconstituted in living *Escherichia coli* cells that we recently constructed (cellulose-synthesizing *E. coli*, CESEC). We demonstrated that inactivating mutations at an important amino acid residue reduced cellulose production. In this study, an interesting loss-of-function mutation occurred on Cys308, whose main chain carbonyl plays an important role for locating the cellulose terminus. Mutating this cysteine to serine, thus changing sulfur to oxygen in the side chain, abolished cellulose production in addition to other apparent detrimental mutations. This unexpected result highlights that the thiol side-chain of this cysteine plays an active role in catalysis, and additional mutation experiments indicated that the sulfur–arene interaction around Cys308 is a key in cellulose-synthesizing activity. Data obtained by CESEC shed light on the function of cellulose synthase in living cells, and will deepen our understanding of the mechanism of cellulose synthase.

1. Introduction

Cellulose is a structural polysaccharide produced by plants, algae, protists, bacteria, and animal tunicates, and one of the most abundant biological polymers on Earth. Cellulose biosynthesis is actually a two-step reaction of (i) polymerization of β1→4-glucan and (ii) crystallization of the resulting polymers (microfibril formation). The enzyme responsible for cellulose biosynthesis is cellulose synthase, which is embedded in the lipid bilayer of cell membranes and is believed to be a heterosubunit complex. Therefore, it has been a difficult protein to analyze, and our biochemical understanding of cellulose biosynthesis remains limited.

The gene encoding the catalytic subunit of cellulose synthase is *cesA* [1,2]. The CesA protein is a glycosyltransferase of the GT-2 family according to the CAZy database. Other polysaccharide synthases are present in the GT-2 family including chitin synthase, β1→3-D-glucan synthase, mannan synthase, and hyaluronan synthase. Hydrophobic cluster analysis of amino acid sequences in an earlier study clarified that these β-glycosyltransferases have two domains (domains A and B), which contain D,D and D,QxxRW sequences, respectively [3]. The function of these highly conserved motif sequences was discussed in that study, for example, domain B was proposed to be important for chain elongation since it was found only in polysaccharide synthases. However, only a few studies [4–6] directly showed that these amino acid residues are involved in polysaccharide synthesis.

Recently, the X-ray crystallographic structure of cellulose synthase was reported from the minimally required subunits of bacterial cellulose synthase, CesA and CesB, of the purple bacterium *Rhodobacter sphaeroides* [7–9]; hereafter, we use the terminology proposed by Delmer (CesA) [1], instead of the original name (BcsA),

Abbreviations: DGC, diguanylate cyclase; c-di-GMP, cyclic-di-guanosine monophosphate; IPTG, isopropyl-β-D-thiogalactopyranoside; FTIR, Fourier-transformed infrared spectroscopy; ATR, attenuated total reflection; PCR, polymerase chain reaction; DP, degree of polymerization.

* Corresponding author.
E-mail address: timai@rish.kyoto-u.ac.jp (T. Imai).
[1] Faculty of Agriculture, Tokyo University of Agriculture and Technology, Fuchu, Tokyo 183-8509, Japan.
[2] Graduate School of Agriculture, Kyoto University, Sakyo, Kyoto 611-0011, Japan.
[3] Department of Plant & Environmental New Resources, College of Life Sciences, Kyung Hee University, 1, Seocheon-dong, Giheung-ku, Yongin-si, Gyeonggi-do 446-701, Republic of Korea.

http://dx.doi.org/10.1016/j.carres.2016.08.009

which was also used other recent studies [7–9]. This is the first three-dimensional structure determination, not only of cellulose synthase, but also for a GT-2 enzyme carrying domain B. This structural model clearly indicates that the two aforementioned aspartic acid residues in domain A are involved in the interaction with the UDP (uridine diphosphate) moiety of the substrate UDP-glucose, and the aspartic acid in domain B functions as the catalytic base in the glucosyltransfer reaction. The famous QxxRW motif in domain B, located in an interfacial α-helix (IF2) running on the inner surface of the lipid bilayer, plays an important role for keeping acceptor cellulose molecules in the correct position for the glucosyltransfer. In addition to these well-known motifs, the structural model of cellulose synthase indicated other important motifs in CesA, such as QTPH and FFCGS. Both of these are close to the acceptor cellulose chains and are proposed to play an important role in glucosyltransfer or chain translocation.

Despite such outstanding structural models that tell us new information about the mechanism of cellulose synthase, it is still important to test the molecular mechanism of cellulose synthase deduced from the X-ray crystallographic structure by observing its enzymatic activity. Site-directed mutagenesis is a clear way to evaluate the role of each amino acid residue in cellulose synthase. Although several previous studies have surveyed the effect of mutating the D,D,D,QxxRW motif in cellulose synthase [5] and chitin synthase [4,6], the mutations in these studies were actually designed based on amino acid sequence alignment [3].

In this study, we designed site-directed mutagenesis of the bacterial CesA protein with the aid of the recently published structural model of the RsCesAB complex [7–9]. Then, we determined the cellulose-synthesizing activity for each of these mutants by expressing bacterial cellulose synthase in recombinant *Escherichia coli*, which is designated as "CESEC" (cellulose-synthesizing *E. coli*) [10]. Our results were consistent with these structural studies, and further suggest that the sulfur–arene interaction around the cysteine residue in the FFCGS motif is key for CesA protein function.

2. Experimental procedures

2.1. Bacterial strains and plasmids

CESEC [10] was used for assaying recombinant cellulose synthase enzymatic activity. Briefly, the *E. coli* strain XL1-Blue was used for expressing cellulose synthase derived from *Gluconacetobacter xylinus* JCM9730 (GxCesA and GxCesB, the minimal requirement for cellulose-synthesizing activity in bacteria). The *gxcesA* and *gxcesB* genes, which are in two sequential open reading frames in the same operon, were inserted into a pQE-80L vector (Qiagen Inc.) for their expression. In addition, DGC (diguanylate cyclase) was expressed together with cellulose synthase. Diguanylate cyclase is an enzyme that synthesizes c-di-GMP (cyclic-di-guanosine monophosphate), an activator of bacterial cellulose synthase [11,12]. The *dgc* gene of *Thermotoga maritima* MSB8, TM1788 [13], was inserted into a pBAD33 vector for expression as a thioredoxin-fused protein. These two vectors were introduced into *E. coli* by chemical transformation, and the transformed *E. coli* were selected using the antibiotics ampicillin and chloramphenicol, resistance to which was induced by the pQE-80L and pBAD33 vectors, respectively.

2.2. Site-directed mutagenesis

A series of site-directed mutagenesis experiments were performed on the *gxcesA* gene in a sub-cloning vector pGEM-T easy (Promega Inc.) using a QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies Inc.), according to the manufacturer's protocol. DNA sequences were verified for the desired mutagenesis, and then the mutated sequence was cut out at the endogenous restriction sites (*EcoR*I sites for D188H, D188N, D189N, and D189Y, and *Bsr*GI and *Sac*II sites for the others). The expression vector containing cellulose synthase (GxCesA and GxCesB) was digested at the same restriction sites for each mutant. Mutated *cesA* DNA fragments were then ligated to the expression vector by using a DNA Ligation kit LONG (Takara Bio Inc., Japan). The ligated plasmid DNA was amplified in *E. coli* HST08 (Takara Bio Inc.) to prepare the expression vector of cellulose synthase containing a particular point mutation. The orientation of the ligated DNA fragment was verified by colony-direct PCR for D188H, D188N, D189N, and D189Y.

2.3. Expression of recombinant proteins

CESEC cells were cultured in 2 × YT medium supplemented with 12.5 μg/mL tetracycline, 100 μg/mL ampicillin, and 50 μg/mL chloramphenicol at 37 °C with orbital shaking at 190 rpm. Protein expression of cellulose synthase and DGC was induced with 0.4 mM IPTG and 0.2% L-arabinose, respectively, when the OD_{600} reached 0.5–0.7. The culture was then maintained with orbital shaking at 28 °C prior to harvesting by centrifugation. The wild type CesA was expressed alongside mutant proteins to compare protein expression and cellulose production for each mutant protein.

2.4. Western blot analysis

Expression of GxCesA and GxCesB was analyzed by SDS-PAGE and western blotting. Cultured cells were collected by centrifugation at 4 °C, and then incubated in SDS-PAGE sample buffer (50 mM Tris-HCl (pH 6.8), 0.5% SDS (sodium dodecyl sulfate), 2% β-mercaptoethanol, 10% glycerol, and 0.01% bromophenol blue) at 4 °C for at least 3 h. Proteins included in the cells in 24 μL culture medium were analyzed with a precast 10–20% polyacrylamide gel (SuperSep Ace, Wako Pure Chemicals Industries Ltd., Japan) alongside a molecular weight marker (Precision Plus Protein All Blue Standards, Bio-Rad Inc.), and transferred onto PVDF membranes (Immobilon-P, Millipore Inc.). The membrane was then incubated with a primary antibody against either GxCesA or GxCesB. The antibodies used were described in our previous studies [10,14] (a polyclonal antibody against a synthetic peptide corresponding to the carboxyl terminus of GxCesA and a loop in the CBD2 domain of GxCesB protein [7], respectively). Membranes were then incubated with anti-rabbit IgG conjugated to horseradish peroxidase (Promega Inc.). Bands of CesA and CesB protein were visualized by a chemiluminescence method to quantitatively evaluate the expression of CesA and CesB proteins. ECL Select (GE Healthcare Inc.) was used for the luminescence reaction and digital images were taken by AE-9300H EZ-Capture MG (ATTO Inc., Japan). The intensity of CesA and CesB bands, which represents the amount of protein in the cells included in 24 μL of culture medium, was determined by ImageJ. Expression of GxCesA and GxCesB protein in each mutant cellulose synthase was evaluated as a percentage relative to wild type, hereafter referred to as the relative expression level.

To verify the correct expression of GxCesA and GxCesB protein, we examined samples prepared by alkaline fractionation [15], which separates the membrane-spanning proteins from other proteins. Briefly, cells treated by lysozyme were solubilized in 0.1 N NaOH solution on ice for 30 min and ultracentrifuged at 100,000 × *g* for 30 min. The pellet, containing membrane proteins, was washed with 5% trichloro acetic acid (TCA), whereas the proteins in the supernatant, including unfolded proteins and soluble proteins, are precipitated by adding TCA in a final concentration of 10%. These precipitated protein samples were air-dried after washing with cold acetone, and then were dissolved in 1% SDS,

50 mM Tris-HCl (pH 8.0), 1 mM EDTA for subsequent SDS-PAGE and western blot analysis.

2.5. Isolation and quantification of synthesized cellulose

Cellulose synthesized by *E. coli* was isolated and quantified as previously described [10]. Briefly, the centrifuged pellet of culture medium was incubated at 100 °C for 30 min in a mixture of glacial acetic acid and concentrated nitric acid (volume ratio of 8:1, hereafter referred to as AN-reagent). The remaining precipitate was used for quantifying cellulose content by the anthrone-sulfuric acid method as previously described [10]. Negative values were treated as zero grams of cellulose. Three independent experiments were performed for each mutant, and statistical tests (Welch's *t*-test) were done by the built-in function of Microsoft Excel (TTEST) to compare cellulose content produced by each construct. Estimated values of the cellulose production per mL of culture in each mutant were normalized to the relative CesA expression level of the mutant to wild type, which was estimated by quantitative western blot analysis as described in *2.4*. Normalized cellulose production was accordingly indicated as a percentage relative to wild type.

2.6. Structural analysis of synthesized cellulose

Cellulose in the culture medium was isolated by chemical washing with AN-reagent, as described in *2.5*. The residue, after repeated washings with water, was analyzed by electron microscopy, electron/X-ray diffraction, and Fourier-transformed infrared spectroscopy (FTIR). Samples were prepared for electron microscopy and electron diffraction on a thin carbon-coated copper mesh. The mesh for electron microscopy was stained with 2% uranyl acetate. A JEM-2000EXII (JEOL Co. Ltd., Japan) was used for both electron microscopy and electron diffraction. Electron micrographs were taken with a CCD camera (MegaView G2; Olympus Soft Imaging Solutions GmbH), and electron diffraction diagrams from a 1-μm diameter region, which was probed by exciting the first condenser lens maximally, were recorded on photographic emulsion films (Kodak SO-163) and developed by Kodak D-19 with full strength by following the manufacturer's protocol. The obtained diffraction patterns were calibrated by a diffraction ring from the (111) plane of gold ($d = 0.2355$ nm).

For X-ray diffraction, washed samples were lyophilized and pressed to make a tablet. X-ray of Ni-filtered Cu Kα generated by a rotating anode (RU-200BH, Rigaku Inc., Japan) operated at 50 kV and 100 mA was collimated with a 0.3 mm-diameter pinhole and directed to the sample tablet. Wide-angle diffraction patterns were recorded on imaging plates (BAS-IP SR 127, Fujifilm Corp., Japan) with a vacuum camera. Camera length was calibrated using NaF powder ($d = 0.23166$ nm).

Spectrum One or Frontier (Perkin Elmer Inc.) was used for FTIR analysis with attenuated total reflection (ATR) attachment. The sample in water was dried on a diamond ATR probe by nitrogen gas, and then spectra were taken with 4 cm^{-1} resolution and 16 integrations in the range of 4000–400 cm^{-1}.

3. Results

3.1. Optimizing conditions for functional analysis of cellulose synthase by CESEC

Prior to analyzing cellulose production by CesA mutants in CESEC, we first determined the optimal duration for expressing cellulose synthase in *E. coli*. The amount of cellulose produced increased for up to 6 h after induction and stayed constant up to 19 h in wild type cellulose synthase (Fig. 1A and Fig. S1). We examined cellulose production at 6 h of induction to evaluate the activity of cellulose synthase over long durations, which is a good indication of a full yield of cellulose production by recombinant GxCesAB in living cells.

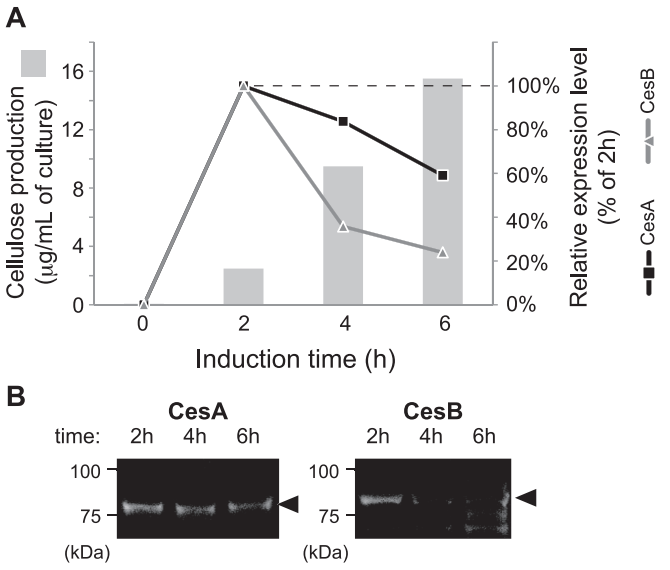

Fig. 1. (A) Accumulation of cellulose produced by CESEC and relative expression level of GxCesA and GxCesB protein over time after inducing protein expression. Protein expression at 2 h of induction is set to 100% for both GxCesA and GxCesB. **(B)** Representative western blot analysis, which was used for plotting expression levels in panel A. Arrowheads indicate the target band.

We also estimated the cellulose production yield normalized to the amount of cellulose synthase expressed. The expression of GxCesA and GxCesB proteins typically reached a maximum at an earlier stage, 2 h after induction, and then gradually attenuated, whereas the synthesized cellulose accumulated up to 6 h after induction as described above (Fig. 1). We analyzed samples at 2 h after induction to correctly normalize cellulose production by cellulose synthase level, given that the relationship between cellulose production and protein expression level at 4 h or later during induction is too complicated to interpret. The level of CesA protein was used for normalizing cellulose production as a percentage to wild type (% of WT) as CesA is a core catalytic subunit of cellulose synthase.

We also analyzed GxCesA and GxCesB protein expressed in *E. coli* by alkaline fractionation [15], which separates membrane proteins as the alkaline-insoluble fraction from the other proteins in the alkaline-soluble fraction. For the mutant proteins tested in this study, western blot analysis revealed that both CesA and CesB proteins were found only in the alkaline-insoluble fraction (Fig. S2), although some mutants showed a very weak signal due to a low expression level. We then concluded that, despite the introduced mutation, GxCesA and GxCesB proteins are correctly expressed in *E. coli* cell membranes.

3.2. Mutation of D,D, and D,QxxRW motifs

Amino acid sequence analysis of CesA indicated the presence of two highly conserved motifs in cellulose synthases and some polysaccharide synthases: the D,D motif and D,QxxRW motif [3]. Recent X-ray crystallographic analyses clearly revealed that these motifs are located in the active site of cellulose synthase [7,8]. Given that these sequences are well conserved in CesA proteins (Fig. S3), we conducted site-directed mutagenesis experiments to inactivate these amino acid residues. The prepared mutant GxCesA proteins were expressed together with wild type GxCesB and DGC.

As expected, most of the inactivating mutants of GxCesA showed significantly reduced cellulose production compared with wild type GxCesA protein as observed by cellulose production at 6 h after induction (full yield), while no significant difference from wild type was found only for the M371A mutant among those tested in this study (Fig. 2A and Table S1, $P = 0.066$ by Welch's *t*-test). Normalized cellulose production also showed that most mutants tested here, except for D189Y and M371A, were clearly inactive (Fig. 2C). We will not discuss the high normalized cellulose production for D189Y as this could be erroneous because of extremely low expression of CesA (Fig. 2B). On the other hand, M371A showed significant cellulose production and enough cellulose was successfully collected for structural analyses as shown in *3.5*. This result is reasonable given that, in this mutation, a less-conserved amino acid (the third residue in the QxxRW motif of CesA, see Fig. S3) was replaced by alanine, which is in general less detrimental for protein structure. Substantial cellulose production was also observed for other mutants in our preliminary data (for example P265A, or *N*/*C*-terminal deletion mutant; data will be shown elsewhere). This clearly means that the cellulose-synthesizing activity in our system can persist when the mutation is not detrimental, and therefore decreased cellulose production in our system is due to decreased activity of the mutant cellulose synthase.

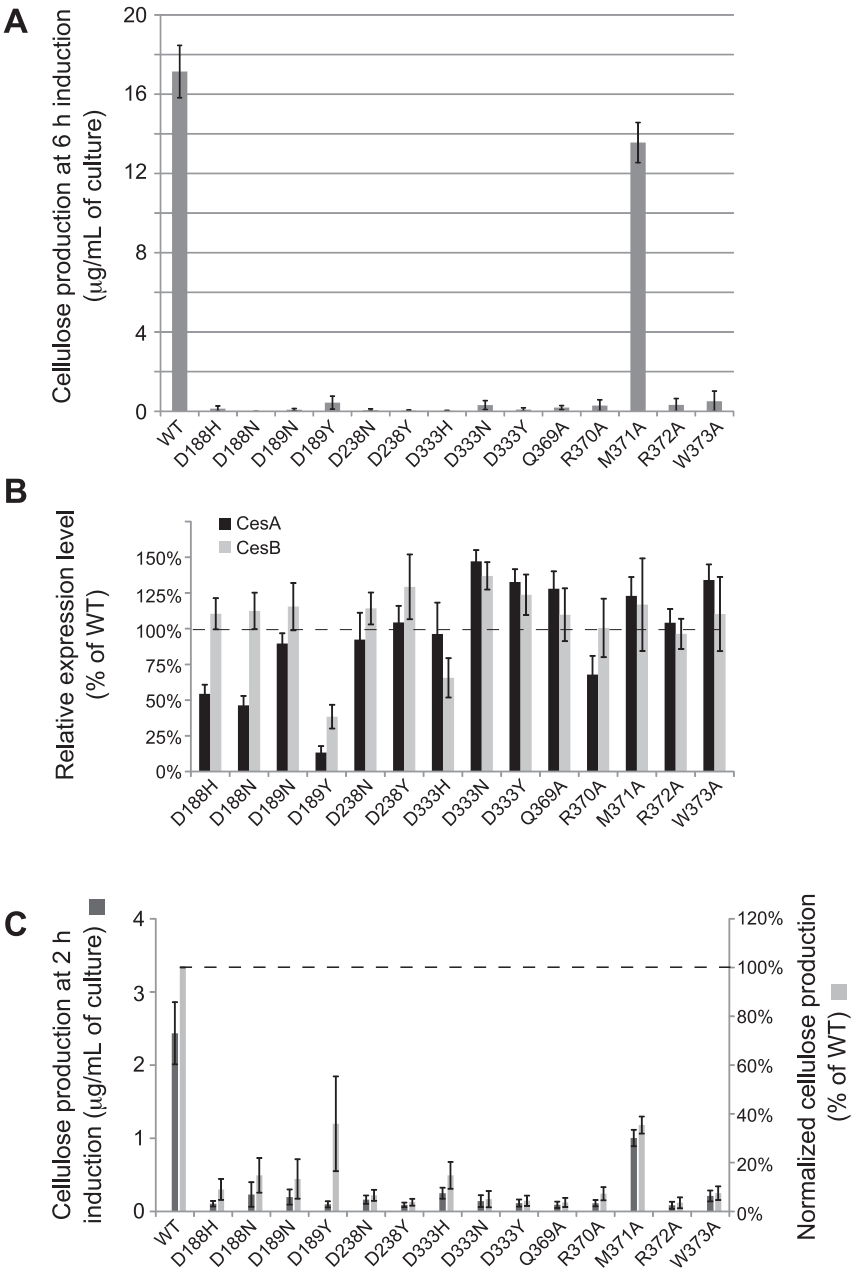

Fig. 2. **(A)** Cellulose production at 6 h of induction (full yield) for wild type and point mutants in the D, D, D, QxxRW motif of GxCesA. Welch's *t*-test showed that all mutants produced significantly lower amounts of cellulose compared to the wild type enzyme ($P < 0.0001$), except the M371A mutant ($P = 0.066$). Cellulose produced by each mutant was determined at $n = 3$, while the determinations for the wild type enzyme were performed at $n = 7$. **(B)** Relative expression level of GxCesA and GxCesB at 2 h of induction. **(C)** Cellulose production at 2 h of induction. Light gray bars indicate the cellulose production yield normalized to CesA expression level, which is shown as the percentage of the wild type value. Error bars indicate the standard error of mean.

3.3. The cysteine side chain in the FFCGS motif is important for function

In addition to the highly conserved sequence motif of D,D,D,QxxRW, the X-ray crystallographic structure newly revealed an important motif sequence in CesA, the FFCGS motif [7]. The cysteine residue in this motif, which lies at the entrance of transmembrane channel, seems to interact with the glucopyranose residue at non-reducing end of the growing chain by using its main chain carbonyl [7]. We tested this hypothesis by mutating this cysteine in the FFCGS motif. Given that the main chain of this cysteine is involved in the proposed interaction, the cysteine could be exchanged to any other amino acid unless largely disturbing the tertiary structure. We then analyzed GxCesA mutants in each of which the corresponding cysteine (C308) was mutated to either of alanine, phenylalanine, histidine, methionine, proline, serine, threonine, or valine.

As shown in Fig. 3A and C and Table S1, all eight GxCesA-C308X mutants tested in this study did not show cellulose production upon expression in CESEC as measured by full yield or normalized production. It is interesting that C308S, in which the side chain sulfur is replaced by oxygen, showed no cellulose production

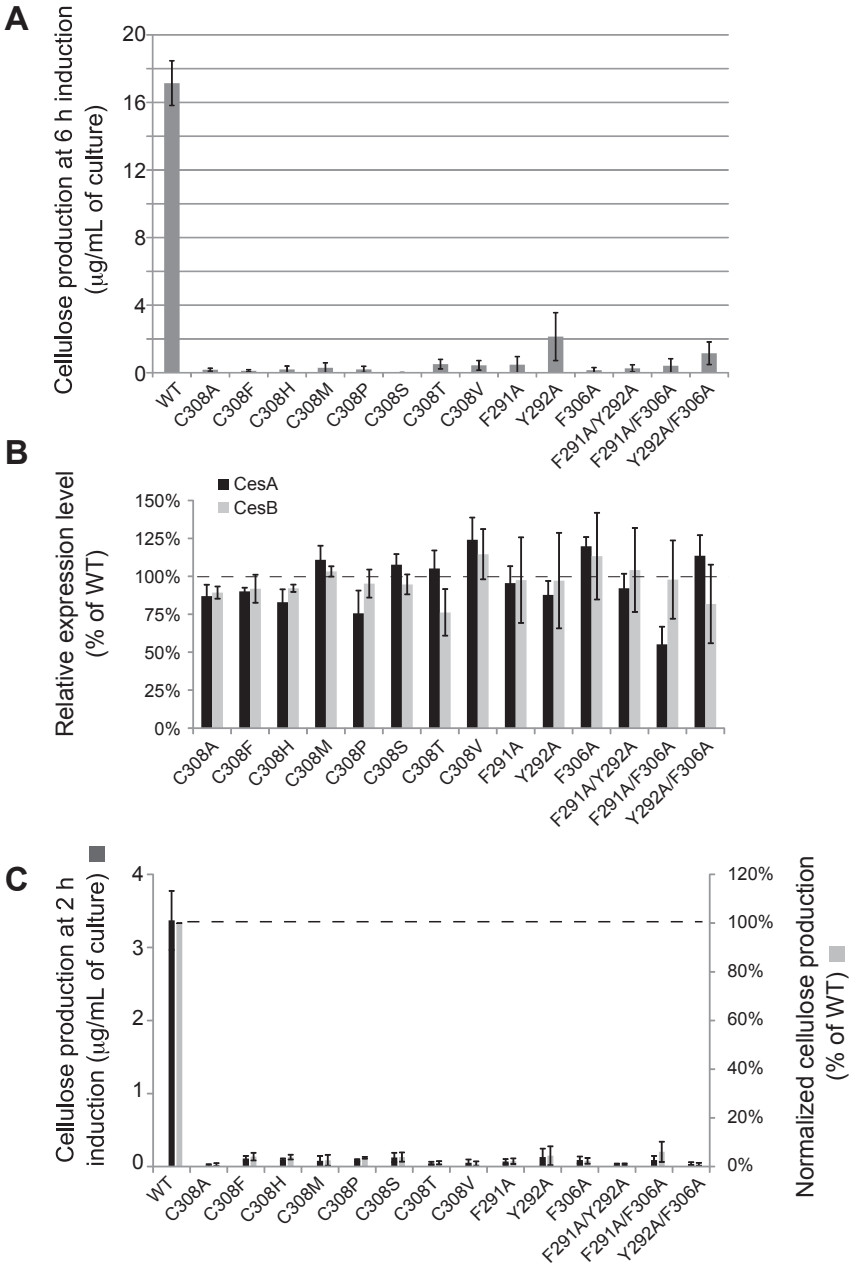

Fig. 3. **(A)** Cellulose production at 6 h of induction (full yield) for the wild type enzyme, point mutants at C308 in the FFCGS motif, and mutations of aromatic residues around this motif (F291, Y292, and F306) to alanine. Welch's *t*-test showed that all mutants produced significantly lower amounts of cellulose compared to the wild type enzyme ($P < 0.0001$). The amount of cellulose produced by each enzyme was quantified at $n = 3$ or 5 (see Table S1), while the amount of cellulose produced by the wild type enzyme was quantified at $n = 7$. **(B)** The relative expression level of GxCesA and GxCesB at 2 h of induction. **(C)** Cellulose production at 2 h of induction. Light gray bars indicate the cellulose production yield of mutants normalized to CesA expression level, which are shown as a percentage of wild type. Error bars indicate the standard error of mean.

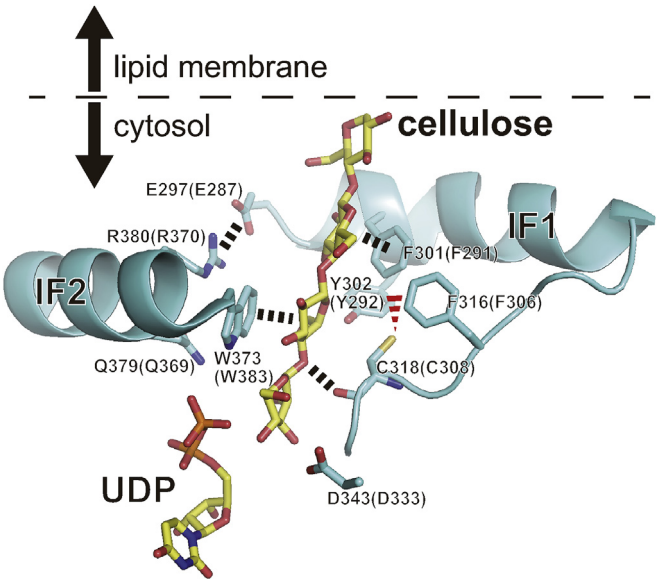

Fig. 4. Close-up of the QxxRW and FFCGS motifs in the structural models of RsCesA in the c-di-GMP–free/UDP–bound state (PDB ID: 4HG6). The cellulose chain and UDP co-crystallized with RsCesA are shown in yellow, and RsCesA protein is shown in cyan. Amino acid numbering is based on the RsCesA protein. Black dashed lines indicate the interaction that Morgan [7] proposed with their crystallographic analysis. The red dashed line indicates the sulfur–arene interaction that we postulated based on our functional study and Morgan's structural analysis. Amino acid residues are numbered with the sequence of CesA protein of *R. sphaeroides* together with that of *G. xylinus* in parentheses, which was based on the alignment in Fig. S3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

despite substantial expression of GxCesA and GxCesB, suggesting the absolute importance of the cysteine side chain in the FFCGS motif. Furthermore, the C308M mutant showed very little activity, indicating that a sulfur in the side chain is not sufficient for activity. These results suggest that the thiol side chain of cysteine in the FFCGS motif is strictly required for the function of cellulose synthase.

3.4. Sulfur–arene interaction may be involved in cellulose synthesis

In order to examine how the cysteine thiol group in the FFCGS motif is involved in cellulose synthase, several other mutants of GxCesA were examined. For all six PDB models deposited (PDB ID: 4HG6, 4P00, 4P02, 5EJ1, 5EIY, and 5EJZ), three aromatic amino acid residues are found around this cysteine (F301, Y302, and F316 for CesA of *R. sphaeroides* (RsCesA), see Figs. 4 and S3). These residues appear to form a pocket surrounding the side chain thiol of cysteine, and the distance between the thiol and the aromatic ring is around 4 Å. An interaction between sulfur and aromatic rings, called a sulfur–arene interaction, is sometimes found in protein structures and protein-ligand interactions [16,17]. To determine if such interactions take place within GxCesA, we performed site-directed mutagenesis on those aromatic amino acid residues (F291, Y292, and F306 in GxCesA, see Fig. S3) systematically changing each to an alanine. Single mutations (F291A, Y292A, and F306A) and double mutations (F291A/Y292A, F291A/F306A, and Y292A/F306A) of GxCesA drastically reduced cellulose production compared to the wild type enzyme (Fig. 3). Therefore, these three aromatic ring amino acids were essential for cellulose synthesis. Finally, no significant cellulose production was observed by GxCesA-C308V and C308F mutants (Fig. 3 and Table S1), indicating that a hydrophobic side chain cannot replace the cysteine residue in the FFCGS motif. All these data support the importance of the cysteine side chain thiol at this residue, and we then propose that a sulfur–arene interaction plays a role to place the cellulose molecular terminal at the correct position for successive glucosyltransfer.

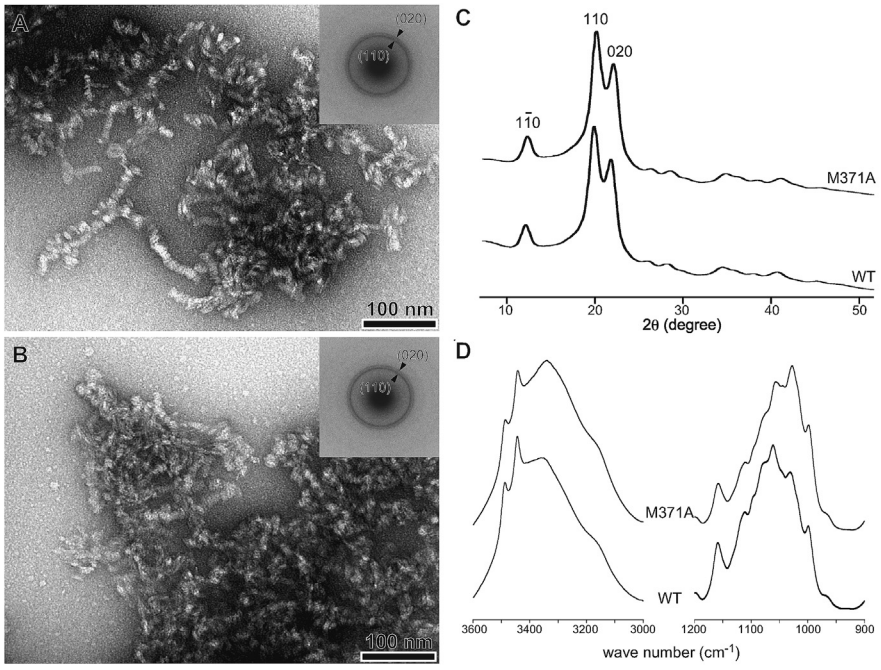

Fig. 5. Structural analysis of cellulose produced by CESEC. **(A and B)** Electron micrographs of the washed product from WT (A) and GxCesA-M371A (B) by negative staining. Insets are electron diffraction diagrams taken from these washed samples: two diffraction rings are derived from the lattice of cellulose II crystal as shown [18]. **(C)** Wide-angle X-ray diffraction profile of cellulose synthesized by recombinant GxCesAB protein with wild type and GxCesA M371A mutant. Major diffraction peaks are indexed by cellulose II as indicated. **(D)** FTIR spectra of cellulose synthesized by recombinant GxCesAB protein with wild type and M371A mutant of GxCesA.

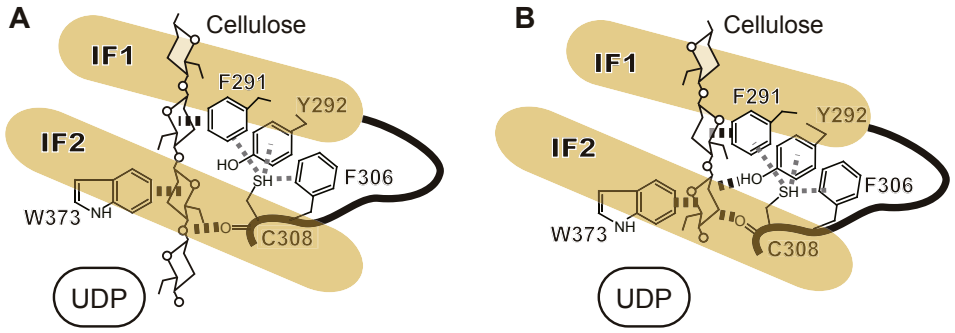

Fig. 6. Schematic diagrams of the putative functional unit composed of the IF1 and IF2 helices and the FFCGS motif in bacterial CesA in two different states. The model in panel A is prepared from the PDB model 4HG6 (c-di-GMP—free and UDP—bound), while panel B is prepared from the PDB model 4P00 (c-di-GMP—bound and UDP—bound). Black and gray dashed lines indicate interactions in CesA proposed by the structural models [7,8] and this functional study (sulfur—arene interaction), respectively. Amino acid residues are numbered with the sequence of CesA protein of *G. xylinus* used in this study.

3.5. Structural analysis of the cellulose produced by a mutant CesA protein

In this study, we found a mutant CesA protein whose activity was not abolished, i.e. M371A. Electron microscopy, electron/X-ray diffraction, and FTIR analysis revealed that the product of this CesA M371A mutant was a globular aggregation crystallized into cellulose II polymorph after chemical washing, as well as that by the wild type (Fig. 5). Electron diffraction diagrams showed two diffraction rings of cellulose II crystal lattices ((110) and (020)) [18], but diffraction from ($1\bar{1}0$) was not confirmed. This rules out that cellulose chains are short enough to let themselves stand perpendicularly on the electron microscopic supporting film, which would allow all these three diffractions to be generated as previously observed in the case of reversing cellulase [19] or cellobiose phosphorylase reaction [20]. Given this result, the synthesized cellulose is probably a polymer long enough to be laid down on the supporting film in both wild type and M371A. These structural characteristics are consistent with our previous studies reported for CESEC [10] and *in vitro* synthesized cellulose by crude enzyme extracts of *G. xylinus* [14].

4. Discussion

4.1. CESEC as a platform for the functional analysis of cellulose synthase

This study functionally analyzed cellulose synthase by using recombinant protein with site-directed mutagenesis, and demonstrated that a point mutation in CesA can alter its enzymatic activity. These mutant proteins were correctly expressed by *E. coli* despite the introduced mutation as shown in Fig. S2. This study putatively concluded that lower cellulose production by mutant cellulose synthases in this study was due to loss of enzyme function.

Given that cellulose synthase reconstituted in living *E. coli* cells was used, this study evaluated enzymatic activity in the cell membrane of a living cell, which is subjected to membrane potential, proton motive force, and solute concentration gradients between the extra- and intracellular space. This is significantly different from *in vitro* assays, which analyze membrane proteins whose extracellular and cytosolic sides are exposed in the same solution environment unless additional techniques are used. This study therefore discusses the structure—function relationship of cellulose synthase in living cells, based on the reported structural models [7–9], and accordingly provides important information in addition to *in vitro* assays, which have the advantage of being amenable to controlling reaction conditions (pH, ligand concentration, temperature, and other parameters).

We however must acknowledge that the reconstituted activity in *E. coli* is not yet able to produce the native cellulose structure (cellulose I microfibril) as shown in Fig. 5 and our previous study [10]. On the other hand, it was shown that the DP (degree of polymerization) of cellulose synthesized by CESEC was as high as 700 [10]. This indicates that the activity reconstituted in *E. coli* cells, which was analyzed in this study, accomplishes the polymerization function but not crystallization. Loss-of-function mutants observed in this study then probably can be described as mutants deficient in polymerization. This is currently the limitation for using CESEC to analyze the enzymatic activity of cellulose synthase.

4.2. A sulfur—arene interaction is involved in cellulose biosynthesis

An unexpected result in this study was that the cysteine residue in the FFCGS motif was not amenable to any mutation examined (i.e., neither alanine, serine, threonine, nor methionine can replace this cysteine and maintain enzyme activity). We then hypothesized that sulfur—arene interactions between the cysteine side chain thiol in the FFCGS motif and the surrounding aromatic side chains of phenylalanine and tyrosine in the IF1 helix are responsible for controlling enzyme activity (Fig. 6). These residues (F301, Y302, F316, and C318 in RsCesA) appear to be located in a similar position in all the six available models [7–9]. This indicates that the IF1 helix and the cysteine in the FFCGS motif operate together even in different states. We then propose that the aromatic residues in the IF1 helix associates with the cysteine side chain in the FFCGS motif by sulfur—arene interaction, which is required for placing the main chain carbonyl of the cysteine residue adequately for the successive glucosyltransfer at the molecular terminus of cellulose. The importance of this sulfur—arene interaction is notably unusual, given that this interaction is not as common as hydrogen bonding, electrostatic interactions, or hydrophobic interactions.

As suggested by Morgan [7], W383 in the QxxRW motif of the IF2 helix appears to interact with the glucopyranose ring of cellulose through CH—π stacking, and Y302 in IF1 seems to make hydrogen bonds with the terminal glucose of growing cellulose chain in the c-di-GMP binding state [8]. These structural models also demonstrate that the terminus of the growing cellulose chain is interposed between the IF1 and IF2 helices [7]. Furthermore, a closer look at the structural model suggests that F301 in the IF1 helix, a counter part of Cys 308 involved in the sulfur—arene interaction proposed in this study, has a CH—π stacking interaction with the glucopyranose ring next to W383 (Fig. 6). These structural observations suggest the importance of IF1 and IF2 helices for CesA

cellulose-synthesizing activity. Given the sulfur–arene interaction between the IF1 helix and the FFCGS motif, it is proposed that IF1, IF2, and FFCGS motifs are functionally connected, and interact with both UDP and cellulose (Fig. 6). As proposed previously [7], the cysteine in FFCGS and the tryptophan in QxxRW play a role in cellulose translocation. Our model supports this hypothesis and furthermore indicates that these two residues operate together in cellulose translocation. The cooperative action of the proposed unit in Fig. 6, which has multiple interactions with UDP-glucose and cellulose, is likely to account for a part of successive glucosyl-transfer mechanism including cellulose translocation, as well as the finger helix movement as shown *in crystallo* [9].

4.3. Introduced mutation could modulate the structure of produced cellulose?

This study showed that GxCesA-M371A still maintains significant activity. However, this is in contrast to a recent report on another GT-2 enzyme, NodC, which synthesizes chitin. The L279A mutant of NodC (an alanine mutation in the third residue of the QxxRW motif), which is equivalent to CesA-M371A in this study, did not exhibit any activity when expressed as a recombinant protein [6]. This implies that mutations in this residue modulate GT-2 enzymes including CesA protein to varying degrees. A closer look at our FTIR data actually revealed that cellulose synthesized by CesA-M371A mutant showed a slightly different IR spectrum in the fingerprint region around 1050 cm^{-1}, as shown in Fig. 5D (the right part of spectra in the figure), indicating that this mutation might modulate the cellulose-polymerizing activity of CesA. Future studies will attempt to explain this difference in cellulose structure for properties such as molecular weight. Accumulating such data with several mutants will shed more light on the mechanism of cellulose synthase.

4.4. Concluding remarks

This study provides a functional analysis of cellulose synthase in living cells. Further studies using this system will deepen our understanding of cellulose biosynthesis. For example, implementing random mutagenesis and the rational design of specific mutations should provide further insight. On the other hand, *in vitro* assays have the advantage of being able to control the concentration of related molecules such as the substrate, activators, or any other ligands or cofactors. Such *in vitro* assays are necessary for analyzing enzyme kinetics, which will provide important insights into the molecular mechanism of cellulose biosynthesis. Given that using CESEC is simple, CESEC will provide complementary information to *in vitro* assays, and will be a powerful tool for studying cellulose biosynthesis.

Acknowledgements

The authors declare no conflicts of interest. This study was supported by CREST, JST to JS, and KAKENHI [25650033; 15H04530] from the Japan Society for the Promotion of Science (JSPS) to TI. Electron microscopic observation was done by the Analysis and Development System for Advanced Materials (ADAM), in RISH, Kyoto University. We would like to thank Editage (www.editage.jp) for English language editing.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.carres.2016.08.009.

References

[1] D.P. Delmer, Ann. Rev. Plant Biol. 50 (1999) 245–276.
[2] C. Somerville, Ann. Rev. Cell Dev. Biol. 22 (2006) 53–78.
[3] I.M. Saxena, R.M. Brown Jr., M. Fevre, R.A. Geremia, B. Henrissat, J. Bacteriol. 177 (1995) 1419–1424.
[4] S. Nagahashi, M. Sudoh, N. Ono, R. Sawada, E. Yamaguchi, Y. Uchida, T. Mio, M. Takagi, M. Arisawa, H. Yamada-Okabe, J. Biol. Chem. 270 (1995) 13961–13967.
[5] I.M. Saxena, R.M. Brown Jr., Cellulose 4 (1997) 33–49.
[6] H.C. Dorfmueller, A.T. Ferenbach, V.S. Borodkin, D.M.F. Van Aalten, J. Biol. Chem. 289 (2014) 23020–23028.
[7] J.L.W. Morgan, J. Strumillo, J. Zimmer, Nature 493 (2013) 181–186.
[8] J.L.W. Morgan, J.T. McNamara, J. Zimmer, Nat. Struct. Mol. Biol. 21 (2014) 489–496.
[9] J.L. Morgan, J.T. McNamara, M. Fischer, J. Rich, H.M. Chen, S.G. Withers, J. Zimmer, Nature 531 (2016) 329–334.
[10] T. Imai, S. -j. Sun, Y. Horikawa, M. Wada, J. Sugiyama, Biomacromolecules 15 (2014) 4206–4213.
[11] P. Ross, H. Weinhouse, Y. Aloni, Nature 325 (1987) 279–281.
[12] O. Omadjela, A. Narahari, J. Strumillo, H. Mélida, O. Mazur, V. Bulone, J. Zimmer, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 17856–17861.
[13] F. Rao, S. Pasunooti, Y. Ng, W. Zhuo, L. Lim, A.W. Liu, Z.-X. Liang, Anal. Biochem. 389 (2009) 138–142.
[14] A. Hashimoto, K. Shimono, Y. Horikawa, T. Ichikawa, M. Wada, T. Imai, J. Sugiyama, Carbohydr. Res. 346 (2011) 2760–2768.
[15] K. Ito, Y. Akiyama, Mol. Microbiol. 5 (1991) 2243–2253.
[16] G. Duan, V.H. Smith, D.F. Weaver, Mol. Phys. 99 (2001) 1689–1699.
[17] L.M. Salonen, M. Ellermann, F. Diederich, Ang. Chem. - Intern. Ed. 50 (2011) 4808–4842.
[18] P. Langan, Y. Nishiyama, H. Chanzy, Biomacromolecules 2 (2001) 410–416.
[19] S. Kobayashi, L.J. Hobson, J. Sakamoto, S. Kimura, J. Sugiyama, T. Imai, T. Itoh, Biomacromolecules 1 (2000) 168–173.
[20] M. Hiraishi, K. Igarashi, S. Kimura, M. Wada, M. Kitaoka, M. Samejima, Carbohydr. Res. 344 (2009) 2468–2473.